

PCT

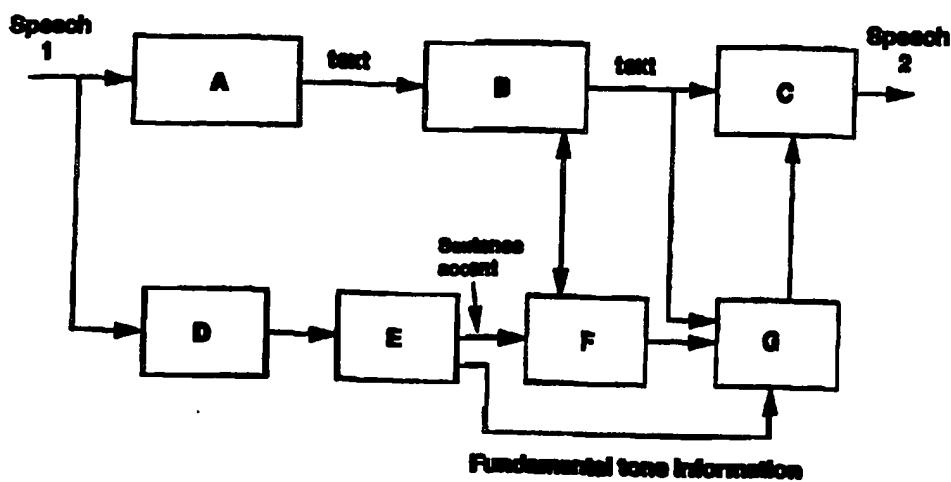
WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : G10L 5/04, G06F 17/28, G10L 5/00		A1	(11) International Publication Number: WO 97/34292
			(43) International Publication Date: 18 September 1997 (18.09.97)
(21) International Application Number: PCT/SE97/00205		(81) Designated States: JP, NO, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(22) International Filing Date: 11 February 1997 (11.02.97)		Published With international search report.	
(30) Priority Data: 9600959-2 13 March 1996 (13.03.96) SE			
(71) Applicant (for all designated States except US): TELIA AB [SE/SE]; Mårbackagatan 11, S-123 86 Farsta (SE).			
(72) Inventor; and			
(75) Inventor/Applicant (for US only): LYBERG, Bertil [SE/SE]; Pingstvägen 5, S-610 72 Vagnhärad (SE).			
(74) Agent: KARLSSON, Berne; Telia Research AB, Rudsjöterrassen 2, S-136 80 Haninge (SE).			

(54) Title: METHOD AND DEVICE AT SPEECH-TO-SPEECH TRANSLATION



(57) Abstract

The present invention relates to a method and device at speech-to-speech translation. A given speech in a first language is recognized in a speech recognition equipment (A). The speech recognition equipment produces a text which is transferred to a translator (B) for translation into a second language. Parallel to these procedures, fundamental tone information is assembled for the first speech. The fundamental tone information influences the prosody generation (G), which influences a text-to-speech converter (C). From the text-to-speech converter a speech in a second language is obtained, the synthesis of which is essentially in accordance with the synthesis of the first speech.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LJ	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

TITLE OF THE INVENTION:

Method and device at speech-to-speech translation.

5 TECHNICAL FIELD

The present invention relates to, from a given natural speech, to produce a corresponding speech in a second language. The second language is produced artificially.

10 PRIOR ART

Attempts to translate between different languages have previously been made. For instance there exist devices which from a given text translate between different languages. Different interpretations of a text however can occur which makes the translator's work more difficult. Other examples of translation are from a speech to another in different languages. In this case the complexity is higher because recognition of the first language is a difficulty in itself. More difficulties will arise if the translated speech shall be reproduced with the voice and characteristics which characterizes the original speaker.

In patent document 9301596-4 a device for improved understanding of speech at artificial translation from one language into another is described. The invention includes an analyzing unit which analyses the duration and the fundamental tone in the speech in the first language. A prosody interpreting unit determines, on basis of the analysis and the information regarding the characteristics of the language, prosody characteristic information in the first language which is used by a prosody generating unit

f r th s cond language f r control of the speech synthesis. A speech synthesis device accordingly effects stress s in the in the sec nd language translated speech which from linguistic point of view correspond to stresses 5 in the first language.

DESCRIPTION OF THE INVENTION

TECHNICAL PROBLEM

- 10 At translation of speech between different languages there is a wish that the characteristics of the speech in the first language is transferred to the second language at the translation. These characteristics are of vital importance for the identification of the speaker of the produced 15 speech. If characteristics are lacking, the produced speech can on one hand be difficult to understand, and on the other give different signals in the speech respective in the characteristics of the speech. The prosodic information content of the speech shall consequently be possible to 20 transfer with principally maintained meaning. Further, there is a wish that the voice of the original speaker shall be reproduced in a lifelike way in the second language.
- 25 Further, there is need to find methods and devices which can be used at direct translation between conversing persons. This can for instance relate to persons who are communicating over a telecommunications network. Other fields which need translations are for instance persons in 30 authority, physicians etc who shall communicate with immigrants in different situations. Especially if the

person with whom the communication is made, speaks a less frequent language, or if the language in itself is well known but a dialect which is difficult to understand is utilized, interpretation problems may arise. The supply of interpreters further are limited, so distance interpretation may sometimes be necessary. The interpreter can in such connections lose much information in ways of expression and body language which are of importance for the interpretation.

10

It is further desirable at the translation to obtain a characteristic in the translated speech which corresponds to the speaker's voice and reproduces his/her state of mind. In the devices and methods which are known, the translated speech is represented by an artificial voice, the characteristics of which does not correspond to that of the first speaker. At an artificial voice of a speaker's verbal presentation it is important that the speaker's voice characteristics in all essentials is translated into the second language. The presentation shall at that in translated sentence be correspondent in respective language. The possibilities for real identification for the person with whom one is talking will at that increase exceedingly.

25 The following invention intends to solve said problems.

THE SOLUTION

The present invention relates to a method and device at speech-to-speech translation. A given speech in a first language is recognized in a speech recognition equipment,

30 A. The speech recognition equipment produces a text which

- is transferred to a translator, B, for translation to a second language. Parallelly with these procedures fundamental tone information for the first speech is produced. The fundamental tone information has an effect on 5 the prosody generation, G, which effects a text-to-speech converter, C. From the text-to-speech converter a speech in a second language is obtained, the synthesis of which essentially is in accordance with the synthesis of the first language.
- 10 The device relates to speech-to-speech translation where a first speech is given. The first speech is given in a first language. The given speech is recognized and translated into a second language. The fundamental tone information in the first language is translated to the second language at 15 which the second speech is produced with a pitch and fundamental tone dynamics corresponding to that of the first speech. The at this produced information will at that announce essentially the same message as the original information in the first speech.
- 20 The fundamental tone of the first speech is normalized and its sentence accents are extracted. This information indicates on one hand the characteristics of the speaker regarding speech, and on the other which parts in the speech that are emphasized. The accents further decide 25 which shades of the translation that can be decisive at the interpretation of the speech. The normalization means that the fundamental tone variation of the speech is divided by the fundamental tone declination of the speech. From normalization of the fundamental tone curve, the dynamics 30 of the speech can be gathered. Further, sentence accents in the incoming speech are classified. The location of said

sentence accents in the second language are determined. The sentence accents consequently are translated into the second language at which an accentuation corresponding to that of the first language is obtained.

- 5 The sentence accent information and the fundamental tone information, fundamental tone declination and fundamental tone dynamics are transferred to a prosody generator. In the prosody generator a written translation of the speech is combined with said other information. This information
- 10 is after that utilized at the text-to-speech conversion at which a speech is produced in a pitch of the voice and an intonation in the second language which is well in accordance with the speech the person would have produced in the second language, at which a part of the speaker's
- 15 identity is transferred.

ADVANTAGES

- The present invention allows that a speech produced by a speaker in a first language is presented with the voice characteristics of the speaker. To a listener of the translated speech this means that the experience is that the translated speech is experienced as directly spoken by the first speaker.
- 25 The utilization of the sentence accents of the first speech and translation of these to the second speech further implies that the characteristics of the second speech is preserved, as well as the intonation at the translation. With the present invention consequently an instrument is given where a given speech at translation into a second
- 30

language is given a corresponding characteristic in the second language.

By the invention is given possibility for two persons to talk to each other in their mother tongues. Use of such systems are of current interest at telecommunication, communication physician/patient etc.

DESCRIPTION OF FIGURES

- Fig. 1 shows the invention in the form of a block diagram.
10 Fig. 2 shows a diagram over the fundamental tone variations over the fundamental tone declination.
Fig. 3 shows a curve over the fundamental tone variation divided by the fundamental tone declination.

15

DETAILED EMBODIMENT

In the following the invention is described on the basis of the figures and the terms therein.

Speech recognition equipments are since before well known to the expert within the speech recognition field. The fundamental functions in speech recognition equipments can be found in books as well as in periodicals. A first speech, speech 1, representing speech from a person, is received by a speech recognition equipment, A, which converts the speech into a text string. The speech recognition equipment evaluates different interpretations which can exist with regard to the interpretation of the speech. The selection of the most probable speech can be made in different ways, for instance by calculus of probability, interpretations of previous sequences in the speech, linguistic selection methods etc. The text string

which has been produced in the speech recognition equipment, A, is after that transferred to a translator, B, which translates the given speech to a text string in the second language. In the translator, B, the fundamental characteristics of the second language is added to the speech of the translated speech. The fundamental characteristics consist of normal accents and pitches in the language.

In order to make a translated speech to give the impression that it is produced by the person in question, it is required that the person's voice characteristics is transferred to the second speech. Further is required that the intonation in the first language is translated into the second language to make it possible to preserve the meaning. Information regarding these voice characteristics are obtained by fundamental tone extraction.

Parallelly with the speech recognition in A, the fundamental tone of the speech, speech 1, is extracted in a fundamental tone extractor, D. The fundamental tone is a combination of fundamental tone declination and fundamental tone variation, Fig.2. These components are separated from each other in E. A normalization of the fundamental tone after that takes place. The normalization means that the variation of the fundamental tone is divided by the declination of the fundamental tone, Fig.3. This information indicates the fundamental tone dynamics of the speaker in the first speech. The sentence accents in the first speech is further determined. The information regarding the sentence accents are transferred to a sentence accent translator, F, which also receives information regarding the translation from translator. The

specific sentence accents which have been identified for the first language now are translated into the second language. I.e. the sentence accents are placed in the second language with regard to the characteristics of the second language. The translation of the sentence accents are after that returned to the translator for linguistic control. The linguistic control includes that the accentuations are modified to the use of the second language. The in this way modified text string is after that transferred to a text-to speech-converter, C, and to a prosody converter, G. The prosody converter further receives information from the sentence accent translator, F, and fundamental tone information from E. In the prosody converter a prosody which is adapted to second language after that is generated. The information from the prosody generator, G, is after that transferred to the text-to-speech converter for generation of a speech, speech 2, the synthesis of which essentially corresponds to the synthesis of the first speech.

20

The invention is not restricted to the above as example shown example or parts of the following patent claims but may be subject to modifications within the frame of the idea of invention.

PATENT CLAIMS

1. Method at speech-to-speech translation, where a first speech, representing a first language, is recognized and translated into a speech in a second language,
5 characterized in that the fundamental tone information of the first speech is translated into the second language, and the second speech is produced with a pitch and a fundamental tone dynamics which is in accordance with the first speech.
- 10 2. Method according to patent claim 1, characterized in that the fundamental tone of the first speech is normalized and that the sentence accents of the first speech are extracted.
- 15 3. Method according to patent claim 1 or 2, characterized in that the sentence accents are translated into the second language.
- 20 4. Method according to any of the previous patent claims, characterized in that information regarding the pitch and fundamental tone dynamics of the first speech is transferred to a prosody generator.
- 25 5. Method according to any of the previous patent claims, characterized in that the first speech is transformed to a first text which is translated into a second text in the second language.
- 30 6. Method according to any of the previous patent claims, characterized in that the sentence accent translation influences the prosody presentation which influences the presentation of the second speech.
7. Method according to any of the previous patent claims, characterized in that the fundamental tone dynamics of the incoming voice is given by maximum f

the fundamental tone variation of the first speech, divided by the fundamental tone declination of the first speech where the fundamental tone declination indicates the pitch of the first speech.

5 8. Device at speech-to-speech translation, where a first speech, representing a first language, is recognized and translated into a second speech in a second language, characterized in that the fundamental tone information of the first speech is translated into the
10 second language, at which the second speech is produced with a pitch and a fundamental tone dynamics corresponding to the first language.

15 9. Device according to patent claim 8,
 characterized in that the fundamental tone of the first speech is normalized and that the sentence
 accents are extracted.

20 10. Device according to patent claim 8 or 9,
 characterized in that the sentence accent information from the first speech is translated into the
 second language.

25 11. Device according to any of the patent claims 8-10,
 characterized in that the sentence accent information is arranged to influence the translation from
 the first language into the second language.

30 12. Device according to any of the patent claims 8-11,
 characterized in that the information regarding the pitch and the fundamental tone dynamics of the first speech is transferred to a prosody generator.

35 13. Device according to any of the patent claims 8-12,
 characterized in that the first speech is

transformed to a text in the second language in a translator.

14. Device according to any of the patent claims 8-13,
characterized in that the prosody generator is
5 influenced by the text and the sentence accent translation.

15. Device according to any of the patent claims 8-14,
characterized in that the prosody generator is
arranged to influence a text-to-speech converter which is
arranged to produce the second speech from the text.

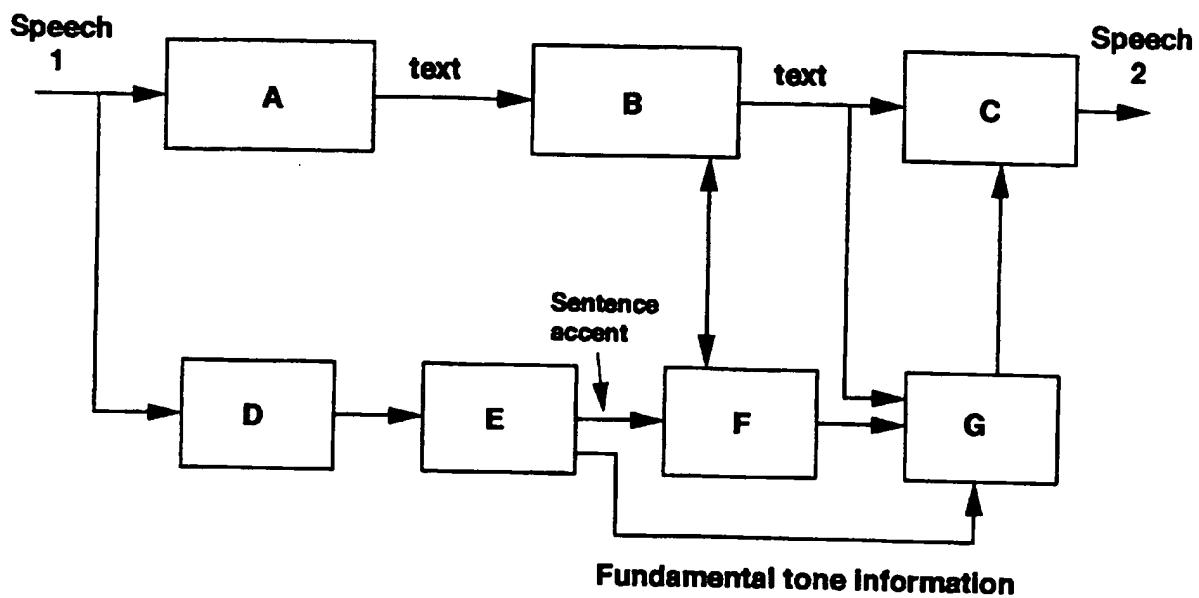


Fig. 1

2/2

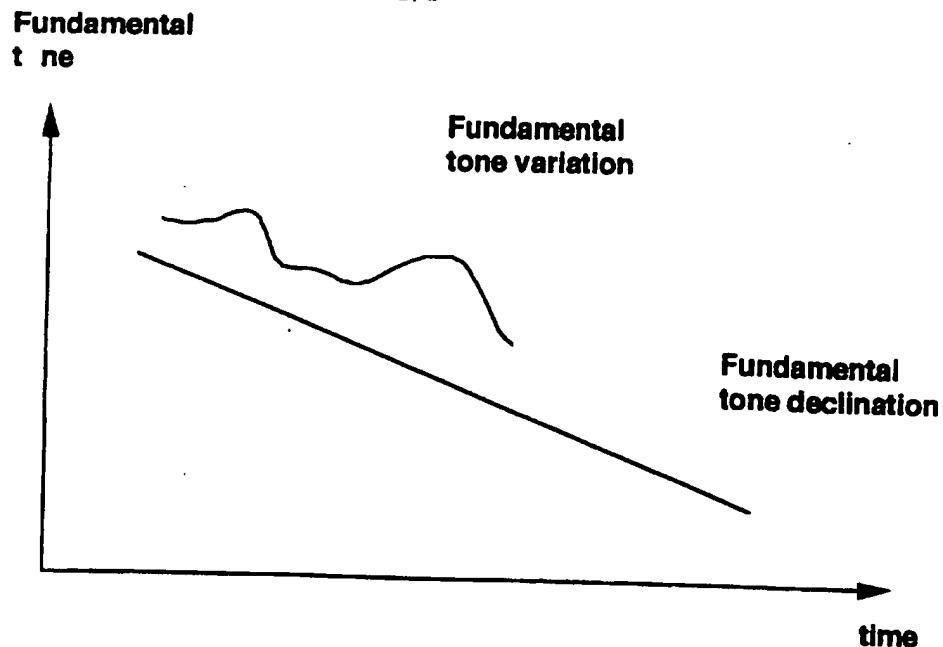


Fig. 2

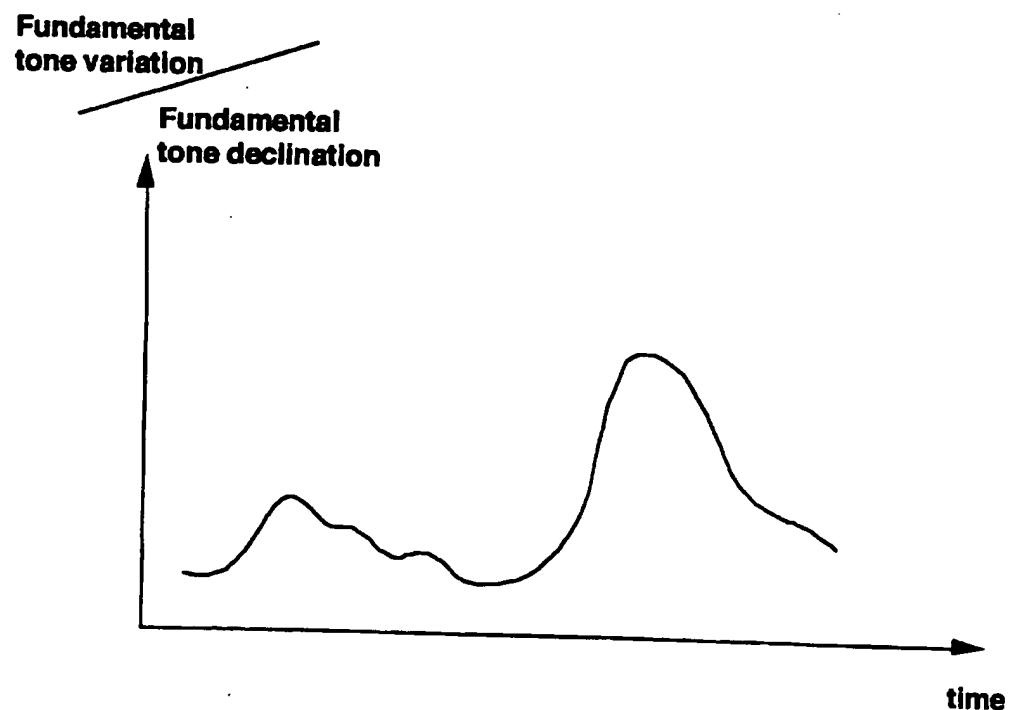


Fig. 3

INTERNATIONAL SEARCH REPORT

1

International application No.

PCT/SE 97/00205

A. CLASSIFICATION OF SUBJECT MATTER

IPC6: G10L 5/04, G06F 17/28, G10L 5/00
 According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC6: G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE, DK, FI, NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPAT, JAPIO, CLAIMS, WPIL, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0624865 A1 (TELIA AB), 17 November 1994 (17.11.94)	1,4-6,8, 12-15
Y	see summary of invention ---	2-3,7,9-11
Y	EP 0664537 A2 (TELIA AB), 26 July 1995 (26.07.95)	2-3,7,9-11
A	see col.3, line 49-55, col.5, line 12 - col.6, line 31, fig.1 --- -----	1,4-6,8, 12-15

<input type="checkbox"/>	Further documents are listed in the continuation of Box C.	<input checked="" type="checkbox"/> See patent family annex.
* Special categories of cited documents:		
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention	
"B" earlier document but published on or after the international filing date	"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone	
"L" document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art	
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family	
"P" document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search	Date of mailing of the international search report
6 June 1997	11 -06- 1997
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Authorized officer Lars Christerson Telephone No. +46 8 782 25 00

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No. PCT/SE 97/00205	
---	--

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP 0624865 A1	17/11/94	JP 6332494 A		02/12/94
		SE 500277 C		24/05/94
		SE 9301596 A		24/05/94
		US 5546500 A		13/08/96
-----	-----	-----	-----	-----
EP 0664537 A2	26/07/95	JP 7181997 A		21/07/95
		SE 9303623 A		04/05/95
-----	-----	-----	-----	-----